CLINICAL PROTEOMIC TECHNOLOGIES FOR CANCER
Advancing Protein Science for Personalized Medicine

eProtein

Letter from the Director

Dear Colleagues,

It is believed that cancer disease processes may have a number of biomarkers associated with them, though the presence of a biomarker by itself may not be useful in clinical practice. To be worthwhile, the biomarker must be a characteristic with a known correlation between the evidential quantity and the disease state that can be measured accurately, easily, and cheaply.

Genomics was the first foray into this promise. The Cancer Genome Atlas (TCGA) and other similar efforts are providing a foundation for defining the genomic alterations in statistically robust numbers of samples from several types of cancer. To achieve the full promise of these rich genomic data sets, an understanding of the functional changes that derive from these genetic alterations is required. Proteomics offers our best hope of translating genetic knowledge into effective biomarkers that can drive the development of new diagnostics and therapeutics for most cancers.

Since its launch, the Clinical Proteomic Technologies for Cancer (CPTC) has achieved significant success in developing an accurate and quantitative protein (and peptide) biomarker assay workflow, incorporating technology development with standards, standard operating procedures (SOPs), data analysis standards, critically needed reagents, and an open access proteomics database. As a result, CPTC has quickly evolved into an international resource that links technologists with cancer biologists and clinical chemists to accelerate the clinical translation of proteomic discoveries.

Looking ahead, CPTC seeks to leverage these advances by conducting proteomics research in a systematic manner for the discovery and development of highly credentialed biomarkers for validation in clinical studies—ultimately launching the next stage of biomedical research.

Empowering and Training New Researchers in Clinical Proteomics

Leaders in New Knowledge - Emerging Technologies

What do underrepresented college students with a strong interest and aptitude in research have in common with the Moffitt Cancer Center's clinical proteomics program? Beginning in January 2010, four undergraduates

continued on page 2

Unique Perspectives from the 510(k) Experience

Mock FDA Premarket Notifications Help Propel Emerging Technologies into Clinical Use

As part of ongoing efforts of the National Cancer Institute (NCI)-FDA Interagency Oncology Task Force subcommittee on molecular diagnostics, members of the NCI's Clinical Proteomic Technology Assessment for Cancer (CPTAC) program submitted two mock 510(k) premarket

continued on page 4

A Clinical Proteomic Technologies for Cancer publication that builds connections throughout the proteomics community

In This Issue

Empowering and Training New Researchers in Clinical Proteomics

Leaders in New Knowledge - Emerging Technologies

(continued from cover)

John Koomen, Ph.D.

Assistant Professor, Molecular Oncology and Experimental Therapeutics
Scientific Director, Proteomics Core Facility
H. Lee Moffitt Cancer Center & Research Institute

began 15 weeks of innovative coursework and laboratory experiments in emerging technologies that they might someday use to individualize treatment for cancer patients.

John Koomen, Ph.D., assistant professor and scientific director of the Proteomics Core Facility, and Cathy Meade, Ph.D., R.N., F.A.A.N., professor and senior member of Health Outcomes and Behavior at H. Lee Moffitt Cancer Center & Research Institute, direct the research training program in clinical proteomics. In discussing the program Leaders in New Knowledge - Emerging Technologies (LINK-ET), Koomen said it is one of many initiatives in clinical proteomics currently emerging at the Moffitt Cancer Center, located on the University of South Florida campus.

> *"Our goal is to build this training initiative into a leadership program and apply quantitative mass spectrometry techniques to clinical specimens using guidelines set by CPTC."*

Not only do the first-of-the-kind curricula enhance Moffitt's training programs, they support emerging technologies with promising clinical applications, and uniquely train college students who typically might encounter traditional barriers to scientific or medical education. Initial program funding was received as an American Recovery and Reinvestment Act (ARRA) supplement to the Cancer Center Support Grant from National Cancer Institute (NCI)'s Center to Reduce Cancer Health Disparities, and supports its strategic objective to impact cancer health disparities through career development. The training program is based on the elements of mentoring, collaboration, and networking; it aims to foster learning, provide positive student/mentor interactions, and cultivate students' interests in cancer research.

Entrance into LINK-ET is restricted to underrepresented students, including ethnic and racial minorities; those from families with low socioeconomic status or from rural areas; individuals who are first in their family to attend college; and students with disabilities. Outstanding candidates possess a grade point average (GPA) above 3.5 and seek a career in cancer research.

Trainees journey through rich curricula in the classroom and laboratory. Topics include, but are not limited to, protein identification, tandem mass spectrometry, proteome cataloging, and protein quantification. Equally rich are experiments accompanying lectures and coursework, and culminating with trainees receiving exposure to cutting-edge technologies such as multiple reaction monitoring (MRM).

Until the program's conclusion in summer 2011, LINK-ET trainees work with a Moffitt scientist-mentor, commit 10 to 15 hours per week to the program

continued on next page

Cathy D. Meade, Ph.D., R.N., F.A.A.N.

Senior Member of Health Outcomes and Behavior, H. Lee Moffitt Cancer Center & Research Institute
Professor, College of Medicine, Department of Oncologic Sciences
University of South Florida

Empowering and Training New Researchers in Clinical Proteomics

Leaders in New Knowledge - Emerging Technologies

(continued from previous page)

(Left to right) Bin Fang, Ph.D., *Staff Scientist*;
Students: Joah Aliancy, Beulah Joseph, Nick Eustace, Karla Schramm;
Danyell Wilson, Ph.D., *Program Coordinator*; John Koomen, Ph.D., *Assistant Professor, Molecular Oncology and Experimental Therapeutics*

during the academic year, and work full-time during the summer. Koomen and Danyell Wilson, Ph.D., a post-doctoral fellow who coordinates the program, expect trainees to apply to graduate or medical schools to become next-generation clinician-scientists. "These outstanding students should be expected to publish their results, continue research from the M.D. or Ph.D. side, and begin their research careers," says Koomen. "Additionally, students will take part in a number of outreach events, such as health fairs and community education programs, to gain an understanding of the relevance of their work in the community, and how it contributes to impacting health disparities," states Meade.

Moffitt received more than $19 million from funds authorized by the 2009 ARRA, including funding to use mass spectrometry to detect/assess multiple myeloma in patient blood and urine. Moffitt's Proteomics Core Facility is among the first to utilize the NCI Clinical Proteomic Technologies for Cancer (CPTC)'s MRM performance mixtures to optimize proteomics platforms in research. Moffitt is now beta-testing reagents (standards kits and reference materials) produced within that program to test and deploy protein/peptide measurement and analysis efforts during LINK-ET training, and also for quality control of shotgun sequencing in the Proteomics Core. The efforts of the CPTC working group have established protocols and benchmarks for performance that can be used to standardize proteomics from laboratory to laboratory.

Koomen emphasized, "These recommendations will be used to improve the function of my laboratory and the Proteomics Core Facility. Our goal is to build this training initiative into a leadership program and apply quantitative mass spectrometry techniques to clinical specimens using guidelines set by CPTC." ■

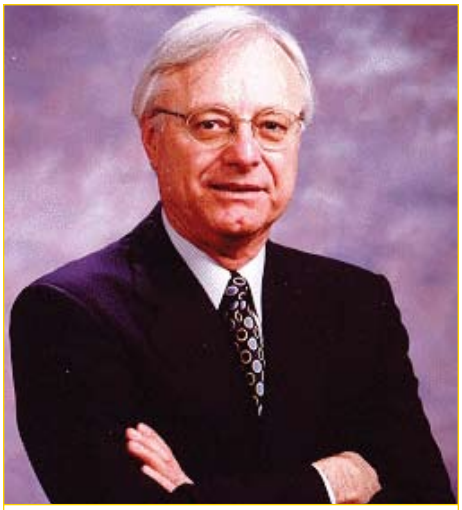
Unique Perspectives from the 510(k) Experience

Mock FDA Premarket Notifications Help Propel Emerging Technologies into Clinical Use

(continued from cover)

submissions for protein-based multiplex assays to the FDA's Office of *In Vitro* Diagnostic Device Evaluation and Safety. Each described a different protein-based platform: an immunological array platform for quantifying glycoprotein isoforms and a multiplex immunoaffinity mass spectrometry (MS) platform for protein quantification. FDA reviewers added mock commentary and asked for additional information or clarification as needed.

The objective was to evaluate analytical measurement criteria and studies required to validate protein-based multiplex assays. Participation by the proteomics research and regulatory communities facilitated an effective, beneficial route to identify analytical issues to address when developing and marketing these emerging technologies.



Fred Regnier, Ph.D.
J. H. Law Distinguished Professor, Analytical Chemistry
Purdue University

N. Leigh Anderson, Ph.D.
Founder, CEO
Plasma Proteome Institute

The mock 510(k) submissions resulted from the 2008 Interagency Oncology Task Force Molecular Diagnostics Workshop, where members of the proteomics and regulatory communities discussed analytical evaluation issues in the development of protein-based multiplex assays for clinical use. The workshop addressed the uncertainty among translational researchers and assay developers about specific analytical measurement criteria needed. The results of the mock 510(k) pre-submissions,[1] as well as workshop issues and recommendations,[2] were published in the February 2010 issue of *Clinical Chemistry*.

Immunological Array

Fred Regnier, Ph.D., J. H. Law distinguished professor of analytical chemistry at Purdue University, Indiana, led the mock 510(k) process for SDIA,

> *"I came away with a huge amount of respect for the FDA…"*

which is a hypothetical immunological array platform that simultaneously measures multiple glycoprotein isoforms in plasma. Regnier noted that he had never tackled a 510(k) submission before this project, but "approached it like a grant proposal; it was similar to writing a proposal about what you thought data looked like." He mentioned that a "huge amount of reading was required, and it was extremely difficult to make up

continued on next page

What is a 510(k)?

According to the U.S. Food and Drug Administration (FDA), a 510(k) is a premarket submission made to FDA to demonstrate that the device to be marketed is at least as safe and effective, that is, substantially equivalent, to a legally marketed device (21 CFR 807.92(a)(3)) that is not subject to PMA (Premarket Approval). Submitters must compare their device to one or more similar legally marketed devices and make and support their substantial equivalency claims.

A 510(k) requires demonstration of substantial equivalence to another legally U.S. marketed device. Substantial equivalence means that the new device is at least as safe and effective as the predicate.[3] This process helps the FDA identify the types of issues it will have to consider when a device based on similar novel technology is submitted for review.

[1] Regnier FE, Skates SJ, Mesri M, et al. Protein-based multiplex assays: mock presubmissions to the US Food and Drug Administration. *Clinical Chemistry.* 2010;56:165-71.

[2] Rodriguez H, Tezak Z, Mesri M, et al. Analytical validation of protein-based multiplex assays: a workshop report by the NCI-FDA Interagency Oncology Task Force on molecular diagnostics. *Clinical Chemistry.* 2010;56:237-43.

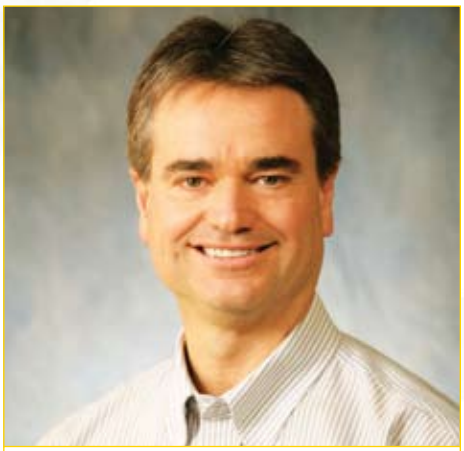
[3] http://www.fda.gov/MedicalDevices/DeviceRegulationandGuidance/HowtoMarketYourDevice/PremarketSubmissions/PremarketNotification510k/default.htm

Unique Perspectives from the 510(k) Experience

Mock FDA Premarket Notifications Help Propel Emerging Technologies into Clinical Use

(continued from previous page)

"The process was a real learning experience—a voyage of discovery."



Scott Patterson, Ph.D.
Executive Director, Medical Sciences
Amgen

AMGEN®

data—especially if one has never performed it—and in a manner that looked real." Of course, researchers are trained to report real data that can be reproduced by others, Regnier cautioned.

What lessons did he learn from the 510(k) experience? "I came away with a huge amount of respect for the FDA, and deep appreciation that they deal with patients who must be treated with their lives at stake, and the product or drugs' proof of intended uses," Regnier emphasized.

Regarding the relevance of the mock 510(k) pre-submission to current technology, Regnier expressed his concern about assaying biologically relevant forms of candidate proteins. He added, "While multiplex isoforms of the same protein can be assayed and we are close to getting individual isoforms, there is no single technique to assess pure isoforms. In the future, one might combine these and other emerging technologies besides antibody- and MS-based ones."

Multiplex Immunoaffinity MS

N. Leigh Anderson, Ph.D., founder and CEO of the Plasma Proteome Institute in Washington, D.C. (http://www.plasmaproteome.org), led the mock 510(k) process for PepCa10, a multiplex diagnostic test using an immunoaffinity multiple reaction monitoring (MRM)-MS platform for protein quantification, which is a hypothetical, MS-based test to measure concentrations of 10 peptides in patient plasma. This technology would ultimately assist physicians in identifying whether a breast biopsy is warranted for low- or high-risk patients. Anderson had no previous regulatory experience. "The process was a real learning experience—a voyage of discovery. Potentially, in a year or so, I might be motivated to do a real 510(k)," said Anderson.

While accustomed to a degree of experimental and statistical rigor, Anderson had not previously faced the level of rigor needed in the premarket submission. "Several iterations of informed commentary from the FDA dealt with very specific aspects of the required performance data, and this extended interaction was very helpful," he keenly observed.

Are regulatory challenges in emerging technologies a high-priority use of funds? Anderson noted that there are multiple stakeholders and issues involved: 1) education of the biomarker community, 2) industry openness to public inspection, and 3) keenness of the FDA to learn about new technologies. "Understanding emerging technologies with their technical differences requires improved, non-classical regulatory focus, and a rigorous science-based approach," said Anderson.

While he doesn't foresee most biomarker researchers investing resources in filing 510(k)s, broader discussion of regulatory issues will be very helpful to those who do. He said one future approach is the possibility of "doing a 510(k) study in a high-end reference lab, such as one of the contract research organizations (CROs) used by pharmaceutical companies. Such studies might be ideal to help move biomarkers into clinical use faster." He suggested that involvement of reference laboratories experienced with assay development on MS platforms might reduce development time by two to five years.

Interestingly, Anderson observed that proteomics researchers, regulatory officials, and the independent laboratory sector were very interested in the mock 510(k)s, specifically in the potential for boilerplate approval processes for mass spectrometry 'home brew' assays, among other things. In discussing clinical use of MS technology, Anderson suggested that a reasonable strategy might encompass bringing diverse groups into the pipeline, including reference labs, instrument and reagent manufacturers, and likely customers.

"In the near future, we might be able to settle on one or two robust, general approaches to regulatory submissions for these MS-based assays, providing a semi-standardized mechanism for people—even

continued on next page

Unique Perspectives from the 510(k) Experience

Mock FDA Premarket Notifications Help Propel Emerging Technologies into Clinical Use

(continued from previous page)

those outside the major diagnostics companies or lacking previous regulatory experience—to do the premarket submissions to the FDA," said Anderson.

Industry Perspective

Scott Patterson, Ph.D., executive director of medical sciences for Amgen, highlighted the industry's roles and interactions with the FDA that could help bring technology to the clinic. "For the diagnostic industry, embracing this new technology is obviously key to its implementation. For the therapeutics industry, exploration of a broader range of protein analytes as predictive biomarkers would no doubt be undertaken if a common assay methodology and attendant instrumentation and software platforms were available that had clear paths to regulatory approval [e.g., those for verification studies and subsequently diagnostics testing]," noted Patterson.

Potential companion diagnostics must also meet the goal of benefiting patients in the global arena—not just one regulatory jurisdiction. Patterson believes that for an MRM assay, the FDA's primary requirement is the device's intended use. "The risk-based classification system that is applied to diagnostics is well thought out and has been applied to a range of diagnostics. The FDA's comments reflect informing those who submitted the document of its shortcomings and how it could be improved to meet the FDA's submission requirements. Primary to such an approach is a clear definition of the intended use of the ultimate device—an area that researchers do not normally consider but need to for diagnostics to become a reality," he explained. The FDA raised specific concerns around appropriate digestion controls, interfering substances, analytical performance evaluation, and instrumentation and software manufacture.

"The engagement of the FDA to review these mock pre-submissions is a great step forward in beginning the dialogue as to what is required for regulatory approval. Few scientists involved in proteomics are familiar with the regulatory approaches for assay evaluation, study design, and statistical analysis to determine the clinical utility (or not) of the candidate biomarker for a defined intended use, including the requirements for all instrumentation and software used in the process. The authors should be encouraged to resubmit the document addressing the concerns of the FDA and repeating this until a satisfactory version has been achieved as a model. This process should involve workshop discussion with all stakeholders—including physicians, payers [e.g., technology assessment center staff], and regulatory bodies—as well as publication of the final product as it stands at that stage," reflected Patterson.

In a larger sense and from the industry perspective, Patterson believes the importance of reproducible, quantitative MS is paramount. He explained, "Although there is an increasing array of antibodies that can be used to construct an accurate ELISA for a given protein, there is an unmet need for the accurate measurement of many proteins and, in some cases, their post-translational modifications. In the latter case, MS-based methods enable rapid development of quantitative assays at reasonable cost—which is critical to the evaluation of candidate biomarkers being found by proteomic methods. Such assays can then be used to evaluate the original hypothesis and prove (or disprove) it in a larger set of samples to determine the utility of the protein biomarker or biomarkers. The greater the investment in the development of these methodologies, the more they will be utilized and have their value realized. Candidate biomarkers are easily found but few are rigorously evaluated to determine their real value; quantitative MS can potentially address this need."

> *"The engagement of the FDA to review these mock pre-submissions is a great step forward in beginning the dialogue as to what is required for regulatory approval."*

Patterson believes several steps will help advance this technology to the clinic, where it might be used to help patients worldwide through the "development of a common assay methodology, and instrumentation [from digestion to analysis] and software that has a path to regulatory approval. The recent partnership between NCI's CPTC initiative and the American Association for Clinical Chemistry [AACC] is a significant step forward to engage additional, important stakeholders in this process.[4] Engaging physician groups and payers, including their health technology assessment centers, will also be critical." Patterson emphasizes that bringing together these groups in forums with the FDA and other international regulatory bodies will be paramount for success. ■

[4] To be featured in an upcoming issue of *eProtein*.

Improving Measurement Quality for Productive Proteomics

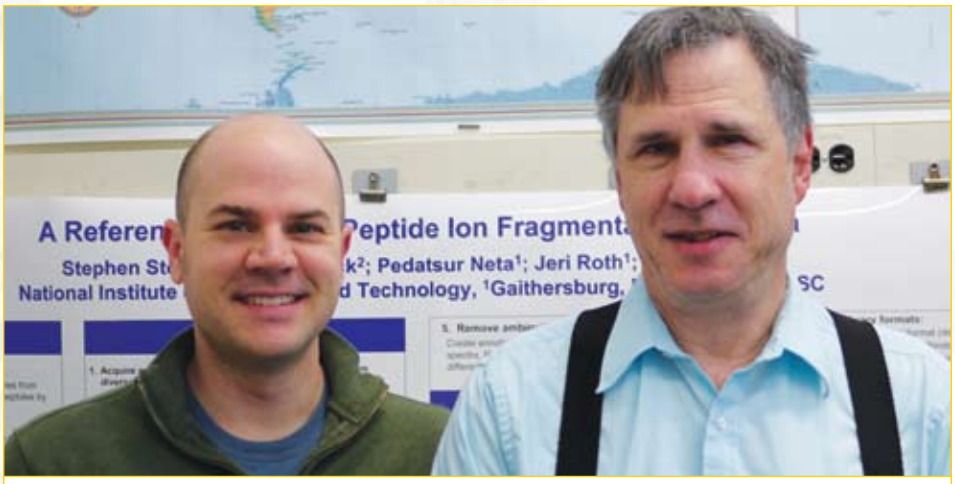
NIST Performance Metrics and MassQC Software

Q. How many liquid chromatography-mass spectrometry (LC-MS) operators does it take to screw in a light bulb?

A. It changes from run to run, machine to machine…

Truth is stranger than fiction. Understanding analytical variability—and its sources—is critical for identifying diagnostic and prognostic protein biomarkers. Before the development of the U.S. National Institute of Standards and Technology (NIST) NISTMSQC software, which implements NIST's specific and quantitative metrics calculations, a major unmet need in liquid chromatography-tandem mass spectrometry (LC-MS/MS)-based proteomics analysis was a tool that assessed system performance quantitatively and evaluated technical variability. Led by Steve Stein, Ph.D., director of the Mass Spectrometry Data Center at NIST, the development of NISTMSQC data-analysis software enabled LC-MS/MS users to monitor and assess run-to-run and/or lab-to-lab variability in key areas of MS-based proteomics experiments.

In a recent publication, collaborator Paul Rudnick, Ph.D., a biologist at NIST, described the 46 system performance metrics that are used by NISTMSQC to analyze data sets from replicate LC-MS/MS analyses. Metrics include chromatographic performance, electrospray source stability, MS1 and MS2 signals, dynamic sampling of ions for MS/MS, and peptide identification.[1] NIST metrics (and NISTMSQC) were developed in collaboration with the National Cancer Institute (NCI)'s Clinical Proteomic Technology Assessment for Cancer (CPTAC) program to evaluate sources of technical variability present in MS-based proteomics experiments. A second major article also used NISTMSQC to describe variability, and to benchmark LC-MS platform performance.[2] More information on NIST is available at http://peptide.nist.gov/ and http://peptide.nist.gov/metrics/.



Paul Rudnick, Ph.D. *(left)*
Biologist, Chemical and Biochemical Reference Data Division
National Institute of Standards and Technology

Stephen Stein, Ph.D. *(right)*
Director, Mass Spectrometry Data Center
National Institute of Standards and Technology

The collaboration between CPTAC and NIST was inspired by the realization that CPTAC had to deal with reproducibility, and NIST was, as Stein recalled, "helping to interpret differing results from different labs." Stein and Rudnick mentioned a major update would be released in summer 2010, with other updates possibly available twice per year. These might add the ability to process data files produced on instruments from other vendors and a graphical user interface (GUI).

Chris Kinsinger, Ph.D., program manager for the NCI's Clinical Proteomic Technologies for Cancer (CPTC) initiative, explained that use of the 46 NIST parameters—and others added continuously—increases the level of robustness in proteomics measurements, creates more awareness of submitted and published data, facilitates and improves measurement quality, and positively impacts mass spectrometry cores at numerous institutions. "While many core facilities might have wondered, 'How well is (my) mass spectrometer performing?' but didn't have the resources or time to analyze this, NIST provided tools to do so," he pointed out. Kinsinger previously worked at NIST performing computer modeling of peptide fragmentation, and later joined CPTC to work on the informatics side of proteomics.

Through their use of NIST metrics, NIST software (NISTMSQC), and commercially available software (e.g., MassQC), community researchers can optimize MS measurement platforms and reduce technical variability that can undermine proteomics research.

continued on next page

[1] Rudnick PA, Clauserb KR, Kilpatrick LE, et al. Performance metrics for liquid chromatography-tandem mass spectrometry systems in proteomics analyses. *Molecular & Cellular Proteomics*. 2010;9:225-41.

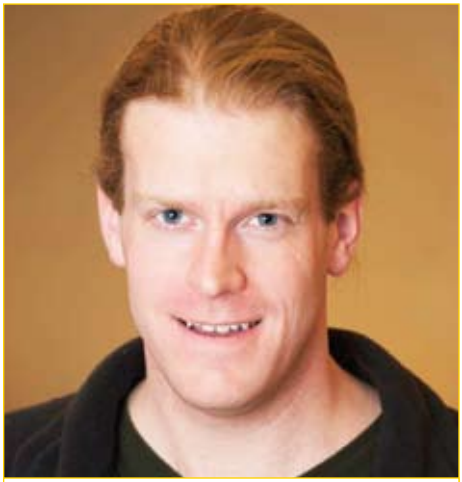
[2] Paulovich AG, Billheimer D, Ham AJ, et al. Interlaboratory study characterizing a yeast performance standard for benchmarking LC-MS platform performance. *Molecular & Cellular Proteomics*. 2010;9:242-54.

Improving Measurement Quality for Productive Proteomics

NIST Performance Metrics and MassQC Software

(continued from previous page)

MassQC was developed by Proteome Software in Portland, Oregon, a for-profit bioinformatics company, with collaboration from NIST (https://www.massqc.com and http://www.proteomesoftware.com). Chris Mason, lead developer of MassQC software, commented upon its commercial release, "The unique design of MassQC makes quality control proteomics more accessible to the LC-MS scientist. It's a powerful and approachable method to get more reproducible experiments." Initially as open-source software, MassQC was first used internally at Proteome Software. Rudnick developed command-line programming and, in 2008, Chris Mason led a team that developed the GUI that still utilized the NIST metrics and algorithms. Beta release began in December 2008 and was followed shortly thereafter by high-profile exposure in meeting workshops and the commercial booth at the 57th American Society for Mass Spectrometry (ASMS) Conference on Mass Spectrometry. Mason noted that this conference "yielded the highest booth traffic ever" and facilitated a core group for beta testing programs.



Chris Mason
Lead Developer, MassQC
Proteome Software

> ***"CPTC is leading the effort to push people to more rigorous protocols. Although it is a challenge, it is important to standardize lab-favorite methods, reagents, and protocols."***

MassQC shows users their current run compared to their historical data for 28 NIST-derived metrics (or parameters). Updates to the software will be issued as needed, primarily based on end-users' responses.

An overarching philosophy is that MassQC "lets people know how their instrument is performing, and that delivers value over time and in the monitoring mode," Mason said. "People understand QC, but they need to monitor. Also, the core lab customers need to show their customers how their instrument is performing by tracking over time and proving instrument performance." Users typically might run multiple standards in between runs, for example, every fourth or fifth run. In essence, the software reduces the time spent assessing the performance of the instrument—while improving reproducibility by examining repeated, standard runs for a small number of known proteins (both complex and simple)—and covers multiple parameters of LC-MS/MS system performance.

According to Mason, "CPTC is leading the effort to push people to more rigorous protocols. Although it is a challenge, it is important to standardize lab-favorite methods, reagents, and protocols. One can then compare data across labs in terms of good science and good data." Mason also noted that labs might critically assess the "value" of the science, and better assess the fiscal resources and commitment needed to maintain a high level of research.

Chris Kinsinger, Ph.D.
Program Manager, NCI CPTC

"Good data sets are generated by the CPTAC community and are run on many instruments. CPTAC uses the same samples that allow users to compare metrics that reflect variations of the machine, not the sample. … A challenge for the community will be to correlate multiple metrics and optimize them to use in a diagnostics role," said Stein.

One end-user, John Klimek of the Oregon Health Sciences University, Portland, routinely utilizes MassQC in the core facility's LTQ and LTQ Velos instruments' workflow (servicing 150 different labs), ranging from proteomics to protein identification. "MassQC offers more robust QC [quality control], points out unknown problems not normally caught (e.g., bad injection times and spray instability, a great metric), and lets you go into troubleshooting faster," said Klimek. He suggested that QC within and between instruments could be utilized, and data might be shared between LTQ instruments in other laboratories.

In terms of clinical proteomics, Klimek offered that increased quality control "gives a way to address traditional objections that proteomics is not reproducible, and also gives a complete picture of how your instrument is working." ■

Researcher Spotlight: *A Korean-American Collaboration*

Youngsoo Kim, Ph.D.
Professor, Department of Biomedical Sciences
Seoul National University College of Medicine

Professor Youngsoo Kim, Ph.D., professor in the Department of Biomedical Sciences at Seoul National University College of Medicine, will soon begin a one-year sabbatical in the laboratory of Amanda Paulovich, M.D., Ph.D., director of the Fred Hutchinson Cancer Research Center's Early Detection Initiative, to focus on quantitative liquid chromatography-multiple reaction monitoring-mass spectrometry (LC-MRM-MS).

Kim recognizes that quantitative LC-MRM-MS protein assays are one of the most effective tools in proteomics research, providing rapid, targeted, multiplexed protein-expression profiling of clinical samples. In fact, due to its inherent capability for higher throughput and multiplexed assays, some believe LC-MRM-MS might complement and/or replace conventional clinical diagnostic immunoassays (e.g., ELISA).

The Paulovich laboratory, one of the major clinical proteomics development sites of the Clinical Proteomic Technology Assessment for Cancer (CPTAC) program, has been developing robust, high-throughput LC-MRM-MS workflows. "Dr. Paulovich's laboratory will be a great place to learn and follow up on this new quantitative proteomics technique," Kim added.

Paulovich and co-investigator Steve Carr, Ph.D., director of Proteomics and Biomarker Discovery at The Broad Institute of MIT and Harvard, recently received $4.8 million in federal stimulus funding from the National Cancer Institute (NCI) to assess the feasibility/scalability of MRM-MS to measure all proteins in the human body. This pilot study was featured in the December 2009 issue of *eProtein*.

The Functional Proteomics Center (FPC), part of the Department of the Ministry of Education, Science and Technology for the Korea Institute of Science and Technology (KIST), is providing approximately $80 million over 10 years to help fund Kim's research in protein biomarker development for diabetes and pancreatic and lung cancers. Kim's sabbatical is in line with the NCI-KIST memorandum of understanding—co-led by Myeong-hee Yu, Ph.D., director of the FPC, and Henry Rodriguez, Ph.D., M.B.A., director of the Clinical Proteomic Technologies for Cancer (CPTC) initiative—which is focused on promoting proteomic technology optimization and standards implementation in large-scale international programs.

Kim recognizes that the benefits of MS-based technology reach far beyond his own research. "Development of protein biomarkers for use in guiding diagnostic and patient treatment decision making will benefit from advances in LC-MRM-MS," he said. "I strongly believe the adaptation of multiplexed assays will be a routine clinical testing element in clinical chemistry," he noted. "Despite sensitivities in the pg/ml range in human blood, non-specific binding is inherent in immunoaffinity assays [ELISAs]. And to ensure that data aren't skewed, costly antibody screening must be performed to increase specificity and sensitivity."

> *"Transition of biomarkers from the bench to the bedside will require more precise, accurate measurement of proteins."*

"Transition of biomarkers from the bench to the bedside will require more precise, accurate measurement of proteins," Kim said. MRM provides excellent specificity for target proteins superior to antibody-based quantitation, and proteins that exist as several hundred copies per cell can be assayed with MRM. "However, improving the limit of sensitivity is needed to measure levels as low as single-digit ng/ml levels in blood," said Kim.

Kim also emphasized the importance of international collaboration in advancing science and technology. "The FPC has interacted closely with both the International Cancer Biomarker Consortium [ICBC], led by Nobel Laureate Dr. Lee Hartwell, and NCI's CPTC initiative, led by Dr. Henry Rodriguez. Since the Paulovich laboratory is actively involved in both programs, my sabbatical will hopefully elevate the collaboration among the FPC, IBCB, and CPTC programs," said Kim. ■

Researcher Spotlight: *A Proteome Genome Public Database (CanProVar)*

Bing Zhang, Ph.D. *(left)*
Assistant Professor, Biomedical Informatics
Vanderbilt University

Jing Li, Ph.D. *(right)*
Lead Author

Bing Zhang, Ph.D., assistant professor of biomedical informatics at Vanderbilt University, conducted his first microarray experiment—and his beginnings in "omics" research—in 2000. In contrast to the one-gene-at-a-time reductionist approach, Zhang remembers collecting data on thousands of genes in just two months. However, due to the lack of appropriate computational tools at the time, data analysis took more than six months. To address this unmet need, Zhang developed and utilized software that was among the first tools for functional interpretation. Today, he estimates that more than 4,000 users have employed data-mining systems in their own research, and more than 300 publications have cited these tools.

Identifying and annotating mutated genes and proteins that are involved in the development and progression of cancer plays a critical role in basic and clinical cancer research. In support of these efforts, Zhang and colleagues developed a human Cancer Proteome Variation Database (CanProVar). CanProVar integrates information on protein sequence variations from six public genome-variation databases and data from two recent, large-scale cancer genome re-sequencing studies, and focuses on cancer-related variations (8,570 specific variations in 2,921 proteins). For ease of use, the authors also developed a user-friendly interface for querying the CanProVar database, accessible at http://bioinfo.vanderbilt.edu/canprovar/.

Zhang and his post-doctoral fellow Jing Li, Ph.D., lead author in the recent article, "CanProVar: A Human Cancer Proteome Variation Database[1]" explain, "CanProVar can serve as a useful tool for the identification of cancer-causative mutations or cancer-biomarkers in the human proteome, the exploration of mutant peptide-based vaccine for cancers, the interpretation of differential peptide expression in shotgun proteomics that are possibly caused by sequence variations, and the explanation of unexpected observations in pull-down experiments for defining protein complexes." Zhang and Li noted that the CanProVar database has been downloaded by scientists from around the world.

In their recent *Human Mutation* article, Zhang and Li report that based on the CanProVar data, cancer-related variations are generally distributed unevenly on chromosomes. Not surprisingly, these variations clustered in cancer-related proteins and were enriched in certain protein domains; such proteins had network interactions with each other. However, the investigators cautioned that potential bias might have been associated with certain data from public and/or study databases (i.e., known cancer genes/hotspots are sequenced more frequently, thus yielding greater numbers of data). Zhang was cautiously optimistic: "I foresee more integrated genomics and shotgun proteomics, and personalized genomics and proteomics. For example, we might generate genomic sequences for individualized patients resulting in proteomic databases."

"Our next step is to improve data annotation, such as standardizing data and vocabulary, and to integrate data sets from different institutions as well as mutations from multiple data sets," said Zhang. Future efforts will also include conducting database research for shotgun proteomics and increasing database size with an enhanced ability to store data, search data sets, and control "false discoveries" with new methods that ensure quality. For example, an analysis based on shotgun proteomics might be validated with genomic sequencing.

Zhang's research group collaborates closely with the Clinical Proteomic Technology Assessment for Cancer (CPTAC) center at the Vanderbilt-Ingram Cancer Center (Daniel Liebler, Ph.D.; director, Jim Ayers Institute for Pre-Cancer Detection and Diagnosis). The focus of this collaboration is on developing and validating tumor tissue-based diagnostics for protein biomarkers that diagnose early colorectal cancer, identifying high-risk disease, and predicting response to therapy. ■

[1] Li J, Duncan DT, Zhang B. CanProVar: a human cancer proteome variation database. *Human Mutation*. 2010:30:1-10.

CLINICAL PROTEOMIC TECHNOLOGIES FOR CANCER

Advancing Protein Science for Personalized Medicine

Upcoming Events

April 17-21, 2010
American Association for Cancer Research (AACR)
101st Annual Meeting
Washington Convention Center
Washington, D.C.

Tuesday, April 20
NCI Meet the Expert Session: Unlocking the Potential of Proteomics Medicine: Answering the What, Why, and How
Remarks by Henry Rodriguez, Ph.D., M.B.A., Program Director, Clinical Proteomic Technologies for Cancer

CPTC 2010 ANNUAL MEETING
Establishing the Standards in Clinical Proteomics
Save the Date September 8-9, 2010 Bethesda, Maryland

Contact Information

For more information about the CTPC, please visit http://proteomics.cancer.gov, or contact us at:

Office of Cancer Clinical Proteomics Research
Center for Strategic Scientific Initiatives
Office of the Director
National Cancer Institute
31 Center Drive, MS 2590
Bethesda, MD 20892
Phone: (301) 451-8883
Email: cancer.proteomics@mail.nih.gov

The NCI Clinical Proteomic Technologies for Cancer seeks to foster the building of an integrated foundation of proteomic technologies, data, reagents and reference materials, and analysis systems to systematically advance the application of protein science to accelerate discovery and clinical research in cancer.

Reagents Data Portal

http://antibodies.cancer.gov
http://dshb.biology.uiowa.edu

Newly Released Antigens and Antibodies

Antigen	Antibody
Fatty Acid Binding Protein, epidermal	CPTC-FABP5-1 CPTC-FABP5-2 CPTC-FABP5-3
Glutathione S Transferase Mu 3	CPTC-GSTMu3-1 CPTC-GSTMu3-2 CPTC-GSTMu3-3
Nucleoside Diphosphate Kinase A	CPTC-NME1-1 CPTC-NME1-1 CPTC-NME1-1
Phosphoserine Aminotransferase 1	CPTC-PSAT-1 CPTC-PSAT-2 CPTC-PSAT-3 CPTC-PSAT-4
Glucose Phosphate Isomerase	CPTC-GPI-1 CPTC-GPI-2 CPTC-GPI-3
Mitogen Activating Protein Kinase 14	CPTC-MAPK14-1 CPTC-MAPK14-1 CPTC-MAPK14-1
Synuclein Gamma	CPTC-SNCG-1 CPTC-SNCG-2

Related News

Argonne National Laboratory is now making labeled and unlabeled 15N proteins available to the scientific community through its website: http://antigens.anl.gov. Many of these proteins were used as immunogens for the antibodies generated for the CPTC reagents program.

Share your *eProtein* story idea or feedback

Help us ensure *eProtein* remains a valuable resource for readers. Email your feedback about a recent issue, or ideas for a future newsletter, to cancer.proteomics@mail.nih.gov.